



DATA ANALYSTE AU SEIN DE LA FAO

Présentation





Projet réalisé par Reda, Paolo et Clément

```
# Nous nous servons de ces librairies  
pour le projet !  
Import pandas as pd  
Import numpy as np  
Import matplotlib.pyplot as plt
```

Etude Guidée 1

Question 1 : Nombre d'humains sur la planète

Dans le fichier CSV concernant la population, une colonne nommée Value correspond au nombres d'habitants dans chaque Pays, nous avons simplement décidé de tout additionner ! Les valeurs sont exprimées par 1000 personnes, donc on viens multiplier par 1000 pour avoir une estimation réele du chiffre

Le résultat que l'on obtient est un chiffre que nous avons trouvé dans un premier temps, quelques peu excessif. Nous avons donc demandé à Google et ChatGPT ce qu'ils en pensaient !

Question 1 : Suite..

Nous avons donc bien un sacré écart entre nos données et les réponses trouvées en ligne, il nous faut donc maintenant identifier d'où provient cet excès de valeurs.

8413993000



7,229 Milliards



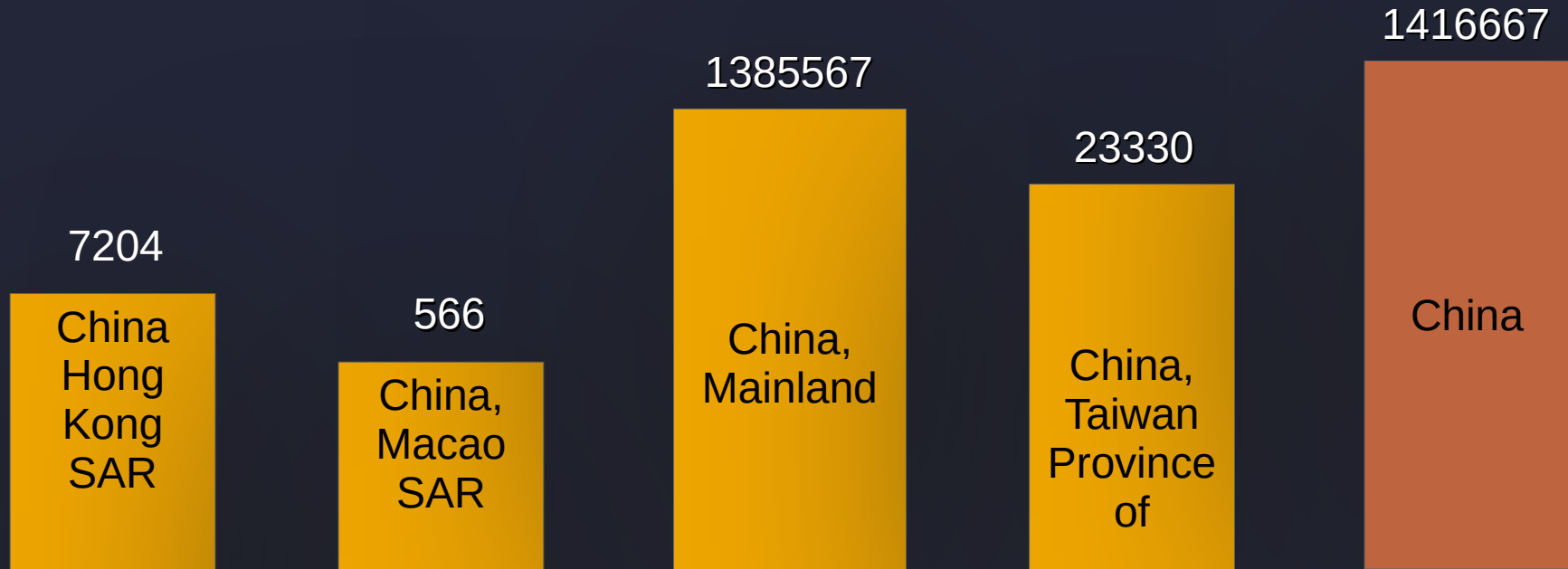
7,1 Milliards



Question 1 : Suite..

On observe que la Chine est comptabilisée deux fois :

- Séparée en Quatre provinces (Hong Kong SAR, Macao SAR, mainland, Taiwan Province of)
- Une Chine qui les regroupe toutes.



Question 1 : Suite..

Si on se penche sur les deux dernières colonnes 'Flag' et 'Flag Description', on remarque une anomalie.

La colonne 'Flag' retourne aucune valeur si ce sont des valeurs officielles. Cependant niveau de la dernière ligne, une 'A' est renseigné dans la colonne, et si on le compare au niveau de la colonne description on observe l'indication : 'Aggregate, may include official, semi-official, estimated or calculated data'

24407		Official data
14539		Official data
14150		Official data
141667	A	Aggregate, may include official, semi-official, estimated or calculated data

Question 1 : Suite..

Après recalcul de la dataframe la population mondiale d'après le dataset est de 6 997 326 000 d'habitant dans le monde en 2013.

Il y a toujours un écart avec les données sur internet, cependant l'écart s'est largement amoindrit et donc le résultat est beaucoup plus précis.

Nous pouvons maintenant procéder à la suite de nos recherches.

Question 2 : Redondances

Ici nous devons identifier les redondances dans un nouveau DataSet, qui liste (les productions et exportations, etc..) tout ce qui est en rapport avec les céréales. Dans notre exemple on utilisera le blé

Ici pour un pays nous identifions plusieurs valeurs mais certaines se complètent. On les identifie depuis la colonne 'Symbole'



Symbole = S

Donnée
standardisée

=

Symbole = Fc

Donnée
calculée

Question 3 : Disponibilité alimentaire (calories, protéines)

Dans cette question il fallait calculer (pour chaque pays et chaque produit) la disponibilité alimentaire en kcal puis en kg de protéines. Avec :

- Population de chaque pays
 - Disponibilité alimentaire donnée pour chaque produit et pour chaque pays en kcal/personne/jour,
 - Disponibilité alimentaire en protéines donnée pour chaque produit et pour chaque pays en g/personne/jour.
-

Nous avons donc procédé en deux étapes

Question 3 : Suite ..

Création d'une table : On assemble les données des tables "animaux" et "végétaux" dans une seule table. Cette table doit inclure la disponibilité alimentaire en kcal/personne/jour et en g/personne/jour pour chaque produit dans chaque pays. Donc on les identifie avec la colonne 'Code élément' ou on viens récupérer la valeur associée.



Calcul de la disponibilité alimentaire totale : Pour chaque pays et chaque produit, on multiplie la disponibilité alimentaire en kcal/personne/jour par la population du pays pour obtenir la disponibilité totale en kcal. Et on fait de même pour la disponibilité en protéines en g/personne/jour, en multipliant par la population pour obtenir la disponibilité totale en kg de protéines.

Question 4 : Ratio énergie/poids

Dans la question précédente, nous avons calculé la disponibilité alimentaire totale en kcal et en kg de protéines pour chaque produit dans chaque pays, en fonction de la population.

Maintenant, pour répondre à la question présente, qui concerne le ratio énergie/poids, nous devons prendre en compte le poids total de la disponibilité alimentaire.

Question 4 : Suite..

Calcul du poids total de la disponibilité alimentaire :
Multiplie la disponibilité alimentaire en kg de protéines (calculée précédemment) par le facteur de conversion approprié pour obtenir le poids total de la disponibilité alimentaire pour chaque produit dans chaque pays.

Calcul du ratio énergie/poids :
Divise la disponibilité totale en kcal (calculée précédemment) par le poids total de la disponibilité alimentaire pour chaque produit dans chaque pays.

Comparaison avec les données disponibles : On vérifie la cohérence de nos résultats en comparant le ratio "énergie/poids" obtenu pour chaque produit avec des données disponibles sur Internet, notre exemple ici : la valeur calorique d'un œuf

Question 5 : Aliments les plus caloriques et protéiques

On identifie ici les 20 aliments avec les ratios énergie/poids les plus élevés, mais cette fois-ci en se concentrant sur les ratios énergie/protéines.

Parmi ces 20, on choisit les 5 aliments qui ont les ratios énergie/protéines les plus élevés.

122 g prot/kg

Poisson
Pelagiques
(Maldives)

87 g prot/kg

Blé
(Tunisie)

81 g prot/kg

Blé
(Azébaïdjan)

84,6 g prot/kg

Blé
(Turkménistan)

84,5 g prot/kg

Blé
(Algérie)

Question 6 : Dispo. intérieure mondiale des végétaux

L'objectif ici est d'obtenir une vision globale de la disponibilité alimentaire des produits végétaux à l'échelle mondiale en termes de calories.

Sélection des produits végétaux :
Identifie tous les produits végétaux dans la table créée précédemment.



Calcul de la disponibilité intérieure mondiale : Pour chaque produit végétal, on somme la disponibilité alimentaire totale en kcal sur tous les pays du monde. Cela donnera la disponibilité intérieure mondiale totale en kcal pour les produits végétaux.

Question 7 : Potentiel alimentaire des végétaux

Notre démarche visait à estimer le potentiel alimentaire des produits végétaux en termes de calories et de protéines, en se basant sur les besoins énergétiques moyens d'un individu.



On a calculé la disponibilité intérieure mondiale des produits végétaux en kcal et en grammes de protéines, puis nous avons comparé ces chiffres aux besoins énergétiques quotidiens moyens d'un individu pour estimer le nombre d'humains pouvant être potentiellement nourris. On exprime le résultat ici en pourcentage

Question 8 : Potentiel alimentaire des végétaux (destinés aux animaux pertes)

Nous arrivons sur une question similaire, nos résultats nous semblent encore peut être pas justes, donc ici, nous allons simplement vous expliquer notre démarche en étapes !

Question 9 : Potentiel alimentaire de la dispo. Mondiale

Question 8/9 : Suite..

Supression de
certaines
colonnes
inutiles de la
dataframe
(df_mergedQ8)
pour simplifier
l'analyse.



Identification des
produits végétaux
destinés à la
nourriture animale
et les pertes de
produits végétaux
grace à la de votre
dataframe
consolidée.



Exclusion de
la
Chine 'Code
élément' 351

Question 8/9 : Suite..

Calcul de la somme des pertes mondiales en termes de calories et de protéines pour les produits végétaux.



Calcul de la somme de la nourriture animale mondiale en termes de calories et de protéines pour les produits végétaux.



Calcul de la somme de la disponibilité alimentaire mondiale en termes de calories et de protéines pour les produits végétaux.

Question 8/9 : Suite..

Somme les valeurs des pertes, de la nourriture animale et de la disponibilité alimentaire mondiale pour estimer le potentiel total en termes de calories et de protéines.



Expression des résultats en pourcentage de la population mondiale, fournissant une perspective sur la portée du potentiel alimentaire des produits végétaux.



Affichage des résultats

Question 10 : Proportion de la sous-nutrition mondiale

Question 11 : Céréales

Question 12 : Sous-nutrition

Question 13 : USA

Question 14 : Thaïlande